

DESCRIPTION

NF- κ B ACTIVATING GENE

This application is a continuation of PCT/JP01/11389 that has an international filing date of December 25, 2001, which designated the United States of America. This application is also a continuation-in-part of Application No. 10/024,298 filed on December 21, 2001, which claims priority under 35 U.S.C. §119(e) on U.S. Provisional Application Nos. 60/258,315 filed on December 28, 2000, 60/278,640 filed on March 26, 2001 and 60/314,385 filed on August 24, 2001. This application also claims priority under 35 U.S.C. §119(a) on Japanese Application Nos. 402288/2000 filed December 28, 2000; 088912/2001 filed on March 26, 2001 and 254018/2001 filed on August 24, 2001. The entire contents of all of the above-identified applications are hereby incorporated by reference.

TECHNICAL FIELD

The present invention relates to a protein capable of activating NF- κ B, a DNA sequence encoding the protein, a method for obtaining the DNA, a recombinant vector containing the DNA, a transformant containing the recombinant vector, and an antibody which specifically reacts with the protein. The present invention also relates to use of the protein, DNA molecule or antibody of the invention in the diagnosis, treatment or prevention of diseases associated with the excessive activation or inhibition of NF- κ B.

The present invention also relates to a method for screening a substance capable of inhibiting or promoting NF- κ B activation by using the protein, DNA, recombinant vector and transformant.

BACKGROUND ART

The transcription factor NF- κ B (nuclear factor kappa B) plays an important role in the transcriptional regulation of various genes involved in inflammation and immunological reactions. NF- κ B is a homo- or heterodimer protein which belongs to the Rel family. In unstimulated conditions, NF- κ B normally resides in the cytoplasm as an inactive form by forming a complex with an I κ B inhibitory protein (Inhibitor of NF- κ B) to mask the nuclear transport signal of NF- κ B.

When cytokines such as interleukin (IL)-1 and tumor necrosis factor (TNF)- α stimulate cells, I κ B is phosphorylated by IKK (I κ B kinase) and degraded by the 26S proteasome through ubiquitination. The released NF- κ B moves to the nucleus, where it binds to the DNA sequence called the NF- κ B binding sequence and induces the transcription of the gene, which is under control of NF- κ B is believed to regulate the expression of genes such as those for immunoglobulins, inflammatory cytokines (e.g., IL-1 and TNF- α), interferons and cell adhesion factors. NF- κ B is involved in inflammation and immune

responses through the expression induction of these genes.

The inhibition of the function or activation of NF- κ B may inhibit the expression of many factors (proteins) involved in inflammatory or immunological diseases or other diseases such as tumor proliferation. Thus, NF- κ B is a promising target for medicaments against diseases caused or characterized by autoimmunity or inflammation [see e.g., Clinical Chemistry 45, 7-17 (1999); J Clin. Pharmacol. 38, 981-993 (1998); Gut 43, 856-860 (1998); The New England Journal of Medicine 366, 1066-1071 (1997); TiPS 46-50 (1997); The FASEB Journal 9, 899-909 (1995); Nature 395, 225-226 (1998); Science 278, 818-819 (1997); Cell 91, 299-302 (1997)].

It is believed that the signal transduction pathway from certain cell stimulation to NF- κ B activation includes many steps mediated by various transmitters such as protein kinases. Therefore it is desirable for more efficient drug discovery to identify the transmitters which play a key role in the pathway, and to focus research on the transmitters to establish a new drug-screening method. Some signaling molecules involved in NF- κ B activation have been identified [e.g., IKK, ubiquitination enzymes and the 26S proteasome described above, as well as TNF receptor associated factor 2 (TRAF2) and NF- κ B inducing kinase (NIK)]. However, most of the mechanism of NF- κ B activation remains unknown, and it has been desired new signaling molecules to be identified and further the NF- κ B activation mechanism to be elucidated.

The object of the present invention is to identify a new gene and protein capable of

directly, or indirectly, activating NF- κ B, and to provide a method of use of them in medicaments, diagnostics and therapy. That is, the present invention provides a new protein capable of activating NF- κ B, a DNA sequence encoding the protein, a recombinant vector containing the DNA, a transformant containing the recombinant vector, a process for producing the protein, an antibody directed against the protein or a peptide fragment thereof, and a process for producing the antibody.

The present invention also provides a method for screening a substance capable of inhibiting or promoting NF- κ B activation, a kit for the screening, a substance capable of inhibiting or promoting NF- κ B activation obtainable by the screening method or the screening kit, a process for producing the substance, a pharmaceutical composition containing a substance capable of inhibiting or promoting NF- κ B activation, etc.

Recently, random analysis of cDNA molecules has been intensively carried out to analyze various genes, which are expressed in vivo. The cDNA fragments thus obtained have been entered for databases and published as ESTs (Expressed Sequence Tags, e.g., <http://www.ncbi.nlm.nih.gov/dbEST>). However, ESTs are merely sequence information, and it is difficult to predict their functions. ESTs are also arranged in UniGene (<http://www.ncbi.nlm.nih.gov/UniGene>), and about 92,000 clusters have been registered until now. However, most of these ESTs have their 5' end nucleotide sequences deleted, and contain no translation initiation site. Therefore it is unlikely that such analysis will directly lead to gene functional analysis such as the analysis of protein functions on the assumption of the determination of mRNA coding regions and the understanding of gene expression control by the analysis of promoters.

On the other hand, one method to elucidate functions of gene products (i.e., proteins) is transient expression cloning method using animal cells [see e.g., "Idenshi Kougaku Handbook (Genetic Engineering Handbook)", an extra issue of "Jikken Igaku (Experimental Medicine)", YODOSHA CO., LTD.]. This method involves transfecting animal cells with a cDNA library constructed using an animal cell expression vector to directly express a functional protein, and identifying and cloning the cDNA based on the biological activity of the protein having an effect on the cells. This method requires no chemical information

(amino acid sequences and molecular weights) regarding the target protein product as a prerequisite, and allows the identification of cDNA clones by detecting specific biological activity of the protein expressed in the cells or culture.

For the efficient expression cloning, there is a need to devise a method of preparing a cDNA library. Several methods have been widely used to construct cDNA libraries [e.g., the method of Gubler-Hoffman: *Gene* 25 (1983); and the method of Okayama-Berg: *Mol. Cell. Biol.* 2 (1982)]. However, most of the cDNA molecules prepared by these methods have their 5' end nucleotide sequences deleted, and thus these methods rarely produce full-length cDNA, a complete DNA copy of mRNA. This is because the reverse transcriptase used to prepare cDNA from mRNA does not necessarily have high efficiency in producing full-length cDNA. Therefore it is necessary to improve these prior art methods in order to efficiently carry out the above expression cloning.

In addition, in order to carry out the functional analysis of genes, it is essential to clone full-length cDNA sequences and express proteins from them. Therefore, it has been necessary to construct cDNA libraries containing enriched full-length cDNA for efficient expression cloning.

The present inventors have intensively studied to solve the above problems. As a result, the present inventors have succeeded in constructing a full-length cDNA library by using the oligo-capping method; establishing a gene function assay system by expression cloning using 293-EBNA cells; and isolating a new DNA (cDNA) encoding a protein having a function of activating NF- κ B by using the assay system. This new DNA molecule induced NF- κ B activation by its expression in 293-EBNA cells. This result shows that this new DNA is a signal transduction molecule involved in NF- κ B activation. Thus, the present invention has been completed.

That is, the present invention relates to:

- (1) A purified protein selected from the group consisting of:
 - (a) a protein that activates NF- κ B which consists of an amino acid sequence represented by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89,

91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178, and 180; and

(b) a protein that activates NF- κ B and consists of an amino acid sequence having at least one amino acid deletion, substitution or addition in an amino acid sequence represented by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180.

(2) A purified protein that activates NF- κ B and comprises an amino acid sequence having at least 50% identity to the protein according to above item (1) over the entire length thereof.

(3) An isolated polynucleotide which comprises a nucleotide sequence encoding a protein selected from the group consisting of:

(a) a protein which comprises an amino acid sequence represented by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180; and

(b) a protein that activates NF- κ B and consists of an amino acid sequence having at least one amino acid deletion, substitution or addition in an amino acid sequence represented by any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180.

(4) An isolated polynucleotide comprising a polynucleotide sequence selected from the

above item (1) or (2) is a membrane protein.

(11) A process for producing a protein comprising,

(a) culturing a transformed cell comprising the isolated polynucleotide according to any one of items (3) to (6) under conditions providing expression of the encoded protein; and

(b) recovering the protein from the culture.

(12) A process for diagnosing a disease or a susceptibility to a disease in a subject related to expression or activity of the protein according to above item (1), (2) or (7) in a subject comprising:

(a) determining the presence or absence of a mutation in the nucleotide sequence encoding said protein in the genome of said subject; and/or

(b) analyzing the amount of expression of said protein in a sample derived from said subject.

In the above-described method, a diagnosis of disease is preferably made when the amount of the protein expressed is 2-fold or higher than normal, or half or less than normal.

(13) A method for screening a compound in respect of activity to inhibit or promote NF- κ B activation, which comprises the steps of:

(a) providing a cell with a gene encoding a protein that activates NF- κ B, and a component that provides a detectable signal associated with activation of NF- κ B;

(b) culturing a transformed cell under conditions, which permit the expression of the gene in the transformed cell;

(c) contacting the transformed cell with one or more compounds;

(d) measuring the detectable signal; and

(e) isolating or identifying an activator compound and/or an inhibitor compound by measuring the detectable signal.

Further, it is preferable to isolate or identify as an activator compound, a compound that increases said detectable signal 2-fold or higher than normal, and to isolate or identify as an inhibitor compound, a compound that decreases said detectable signal half or less than normal.

(14) A process for producing a pharmaceutical composition, which comprises the steps of:

(a) providing a cell with a gene encoding a protein that activates NF- κ B, and a component

capable of providing a detectable signal;

(b) culturing a transformed cell under conditions, which permit the expression of the gene in the transformed cell;

(c) contacting the transformed cell with one or more candidate compounds;

(d) measuring the detectable signal;

(e) isolating or identifying an activator compound and/or an inhibitor compound by measuring the detectable signal; and

(f) optimizing the isolated or identified compound as a pharmaceutical composition.

Further, it is preferable to isolate or identify as an activator compound, a compound that increases said detectable signal 2-fold or higher than normal, and to isolate or identify as an inhibitor compound, a compound that decreases said detectable signal half or less than normal.

(15) A kit for screening a compound in respect of activity to inhibit or promote NF- κ B activation, which comprises:

(a) a cell comprising a gene encoding a protein that activates NF- κ B, and a component that provides a detectable signal upon activation of NF- κ B; and

(b) reagents for measuring the detectable signal.

(16) A monoclonal or polyclonal antibody that specifically binds to the protein according to above item (1), (2) or (7).

(17) A process for producing a monoclonal or polyclonal antibody according to above item that specifically binds to the protein of above item (1),(2) or (7), which comprises administering the protein according to above item (1), (2) or (7) as an antigen or epitope-bearing fragments to a non-human animal.

(18) An antisense oligonucleotide complementary to the polynucleotide according to any one of above items (3) to (6), which prevents NF- κ B activator protein expression.

(19) A ribozyme which inhibits NF- κ B activation by cleavage of RNA that encodes the protein of above item (1), (2) or (7).

(20) A method for treating a disease, which comprises administering to a subject an amount of a compound screened by the process according to above item (13), and/or a

monoclonal or polyclonal antibody according to above item (16), and/or an antisense oligonucleotide according to above item (18), and/or a ribozyme according to above item (19) effective to treat a disease selected from the group consisting of inflammation, autoimmune diseases, infectious diseases, cancers and bone diseases.

(21) A pharmaceutical composition produced according to the process of item (14) as an inhibitor or promoter of NF- κ B activation.

(22) A pharmaceutical composition according to item (21) for the treatment of inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischemic disorders.

(23) A method of treating inflammation, autoimmune diseases, cancers, infectious diseases, bone diseases, AIDS, neurodegenerative diseases, or ischemic disorders, which comprises administering a pharmaceutical composition produced according to the process of above item (14) to a patient suffering from a disease associated with NF- κ B.

(24) A pharmaceutical composition which comprises a monoclonal or polyclonal antibody according to item (16) as an active ingredient.

(25) A pharmaceutical composition which comprises an antisense oligonucleotide according to item (18) as an active ingredient.

(26) The pharmaceutical composition according to item (24) or (25), wherein the target disease is selected from the group consisting of inflammation, autoimmune diseases, infectious diseases, cancers, bone diseases, AIDS, neurodegenerative diseases and ischemic disorders.

(27) A method for obtaining a novel gene having a function, which comprises at least the following steps:

- (a) constructing a full-length cDNA library by the oligo-capping method;
- (b) cotransfecting the full-length cDNA and a plasmid containing a factor emitting a signal indicative of the presence of a protein having the function into cells; and
- (c) selecting a plasmid emitting the signal.

It should be noted that a novel gene having a function according to the present invention refers to, for example, a nucleic acid molecule encoding a protein having biological function.

(30) An insoluble substrate to which polynucleotide comprising all or part of the nucleotide sequences selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88 and 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179, are fixed.

10

The contents of the specifications and/or drawings of Japanese Patent Applications Nos. 2000-402288, 2001-088912 and 2001-254018, and U.S. Provisional Applications Nos. 60/258,315, 60/278,640 and 60/314,385, which from the bases of priority of the instant application, are incorporated herein.

BRIEF DESCRIPTION OF THE DRAWINGS

Fig. 1 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 5) in Example 3, the axis of abscissa is MG-132 concentration and the transversal axis is relative luciferase activity where relative luciferase activity is taken as 100% under conditions of non-addition of MG-132 ($0 \mu\text{M}$). (Relative luciferase activity at various concentrations was divided by relative luciferase activity under conditions of non-addition of MG-132, and expressed as a percentage.)

Fig. 2 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 9) in Example 3.

Fig. 3 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 17) in Example 3.

Fig. 4 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 21) in Example 3.

Fig. 5 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 35) in Example 3.

Fig. 6 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 37) in Example 3.

Fig. 7 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 41) in Example 3.

Fig. 8 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 53) in Example 3.

Fig. 9 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 57) in Example 3.

inhibitor MG-132 (SEQ ID NO: 162) in Example 3.

Fig. 25 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 168) in Example 3.

Fig. 26 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 170) in Example 3.

Fig. 27 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 172) in Example 3.

Fig. 28 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 176) in Example 3.

Fig. 29 is a graph showing NF- κ B reporter activity inhibition by the proteasome inhibitor MG-132 (SEQ ID NO: 178) in Example 3.

Explanation of the Sequence Listing

SEQ ID NO: 181 is a primer.

SEQ ID NO: 182 is a primer.

BEST MODE FOR CARRYING OUT THE INVENTION

At first, in order to further clarify the basic feature of the present invention, the present invention is explained by following how the present invention is completed. In order to obtain a new gene having a function of activating NF- κ B, the following experiments were carried out as shown in the examples. First, using the oligo-capping method, a full-length cDNA was produced from mRNA prepared from normal human lung fibroblasts (purchased from Sanko Junyaku Co., Ltd.), and a full-length cDNA library was constructed in which the cDNA was inserted into the vector pME18S-FL3 (GenBank Accession AB009864). Next, the cDNA library was introduced into E. coli cells, and plasmid preparation was carried out per clone. Then, the pNK κ B-Luc reporter plasmid (STRATAGENE) containing a DNA encoding luciferase under control of a promoter activated by NF- κ B and the above full-length cDNA plasmid were cotransfected into 293-EBNA cells (Invitrogen). After 24 or 48 hours of culture, luciferase activity was measured, and the plasmid with significantly increased

luciferase activity compared to that of a control experiment (vector pME18S-FL3 is introduced into a cell in place of a full-length cDNA) was selected (the selected plasmid showed a 5-fold or more increase in luciferase activity compared to that of the control experiment), and the entire nucleotide sequence of the cDNA cloned into the plasmid was determined. The protein encoded by the cDNA thus obtained shows that this protein is a signal transduction molecule involved in NF- κ B activation.

The present invention is described in detail below.

In the present invention, activation of NF- κ B refers to direct or indirect activation of NF- κ B (including induction of NF- κ B activation) when a gene is introduced into a suitable cell and the protein encoded by the gene is excessively expressed. Activation of NF- κ B can be measured, for example, by an assay using an NF- κ B dependant reporter gene. In the assay, activation may be reflected by increasing the reporter activity compared to control cells (cells into which the vector only was introduced). Increase in reporter activity is preferably by a factor of 1.5 or more, more preferably by a factor of 2 or more, and still more preferably by a factor of 5 or more.

Reporter activity can be measured by cloning a polynucleotide (e.g. cDNA) encoding the protein to be expressed into a suitable expression vector, co-transfecting the expression vector and an NF- κ B dependant reporter plasmid into a suitable cell, and after culturing for a certain period, then measuring reporter activity. Suitable expression vectors are well known to those skilled in the art, examples of which include pME18S-FL3, pcDNA3.1 (Invitrogen). The reporter gene can be one which enables a person skilled in the art to easily detect the expression thereof, and examples include a gene encoding luciferase, chloramphenicol acetyl transferase, or β -galactosidase. Use of a gene encoding luciferase is most preferable, and examples of an NF- κ B dependent reporter plasmid include pNF- κ B-Luc (STRATAGENE). Suitable cells include cells which exhibit an NF- κ B activation response to stimulation by IL-1, TNF- α and the like. Examples include 293-EBNA cells. Cell culture and introduction of genes into cells (transfection) can be performed and optimized by a person skilled in the art by known techniques.

As a preferable method, 293-EBNA cells are inoculated on 5% FBS (Fetal Bovine

Serum) containing DMEM medium (Dulbecco's Modified Eagle Medium) in a 96-well cell culture plate to a final cell density of 1×10^4 cells/well, and cultured for 24 hours at 37°C in the presence of 5% CO_2 . Then, reporter plasmid pNF- κ B-Luc (STRATAGENE) and the expression vector are cotransfected into the cells in a well using FuGENE 6 (Roche). After 24 hours of culture at 37°C , NF- κ B activation is then measured by measuring luciferase activity using a long term luciferase assay system, Picagene LT2.0 (Toyo Ink Mfg). For example, luciferase activity can be measured using PerkinElmer's Wallac ARVOTMST 1420 MULTILABEL COUNTER. The method for gene introduction by FuGENE6, and measurement of luciferase activity by Picagene LT2.0 can be performed respectively according to the attached protocols. In a method of gene introduction with a 96-well plate using FuGENE6, the amount of FuGENE6 per 1 well is suitably 0.3 to $0.5 \mu\text{l}$, preferably $0.3 \mu\text{l}$; the amount of pNF- κ B-Luc plasmid is suitably 50 to 100ng, preferably 50ng; the amount of expression vector is suitably 50-100ng, and preferably 100ng. An ability to activate NF- κ B refers to an ability to increase the reporter activity (luciferase activity) relative to the control experiment (cells into which only a null vector was introduced). Increase in reporter activity is preferably by a factor of 1.5 or more, more preferably by a factor of 2 or more, and still more preferably by a factor of 5 or more.

Related to the amino acid sequences of any one of SEQ ID NOS. 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180, the present invention provides for a protein that:

- (a) comprises one of the above amino acid sequences;
- (b) is a peptide having one of the above amino acid sequences;
- (c) activates NF- κ B and consists of an amino acid sequence having at least one amino acid deletion, substitution or addition in the above amino acid sequences;
- (d) comprises an amino acid sequence, which has at least 95% identity, preferably at least 97-99% identity, to an amino acid sequence of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21,

23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 or 180, over the entire length thereof.

"Identity" as known in the art, is a relationship between two or more protein sequence or two or more polynucleotide sequences, as determined by comparing the sequences. In the art, "identity" also means the degree of sequence relatedness between protein or polynucleotide sequences, as determined by the match between protein or polynucleotide sequences, as the case may be, as determined by the match between strings of such sequences. "Identity" and "similarity" can be readily calculated by known methods. Preferred methods to determine identity are designed to give the largest match between the sequences tested. Methods to determine identity and similarity are codified in publicly available computer programs. "Identity" can be determined by using the BLAST program (for example, Altschul SF, Gish W, Miller W, Myers EW, Lipman DJ., J. Mol. Biol., 215: p403-410 (1990), Altschul SF, Madden TL, Schaffer AA, Zhang Z, Miller W, Lipman DJ., Nucleic Acids Res. 25: p3389-3402 (1997)). Where software such as BLAST is used, it is preferable to use default values. The main initial conditions generally used in a BLAST search are as follows, but are not limited to these.

An amino acid substitution matrix is a matrix numerically representing the degree of analogy of each pairing of each of the 20 types of amino acid, and normally the default matrix of BLOSUM62 is used. The theory of this amino acid substitution matrix is shown in Altschul S.F., J. Mol. Biol. 219: 555-565 (1991), and applicability to DNA sequence comparison is shown on States D. J., Gish W., Altschul S.F., Methods, 3: 66-70 (1991). In this case, optimal gap cost is determined by experience and in the case of BLOSUM62 preferably parameters of Existence 11, Extension 1 are used. The expected value (EXPECT) is the threshold value concerning statistical significance for a match with a database sequence, and the default value is 10.

As one example, a protein having, for example, 95% or more identity to the amino

The Examples described below demonstrate that the protein consisting of an amino acid sequence of any one of the above SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180, is capable of activating NF- κ B.

Related to the polynucleotide sequence of any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179, the present invention further provides an isolated polynucleotide that:

- (a) comprises a nucleotide sequence, which has at least 95% identity, preferably at least 97-99% identity to any one of the above sequences;
- (b) is a polynucleotide of any one of the above sequences; or
- (c) has a nucleotide sequence encoding a protein which has at least 95% identity, preferably, at least 97-99% identity, to the amino acid sequence of any one of SEQ ID NOs: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147,

The polynucleotide of the present invention may be either in the form of a DNA such as cDNA, a genomic DNA obtained by cloning or synthetically produced, or may be in the form of RNA such as mRNA. The polynucleotide may be single-stranded or double-stranded. The double-stranded polynucleotides may be double-stranded DNA, double-stranded RNA or DNA:RNA hybrid. The single-stranded polynucleotide may be sense strand also known as coding strand or antisense strand also known as non-coding strand.

18

addition relative to the protein of any one of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180, and the DNA encoding the protein. The number of mutations is preferably up to 10, more preferably up to 5, most preferably up to 3.

The substitutions of amino acids are preferably conservative substitutions, specific examples of which are substitutions within the following groups: (glycine, alanine), (valine, isoleucine, leucine), (aspartic acid, glutamic acid), (asparagine, glutamine), (serine, threonine), (lysine, arginine) and (phenylalanine, tyrosine).

Based on the nucleotide sequences (e.g., SEQ ID NO: 2) encoding a protein consisting of an amino acid sequence of any one of SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180 or a fragment thereof, those skilled in the art can routinely isolate a DNA with a high sequence similarity to these nucleotide sequences by using hybridization techniques and the like, and obtain proteins having the same NF- κ B activating activity as the protein having of an amino acid sequence of any one of SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180. Thus, the present invention also includes a protein that activates NF- κ B and comprises an amino acid sequence having a high identity to the amino acid sequence of any one of the above SEQ ID NO:1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119,

121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180. "High identity" refers to an amino acid sequence having an identity of at least 90%, preferably 95%, and more preferably at least 97% over the entire length of an amino acid sequence expressed by any one of the above SEQ ID NO: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180. The proteins of the present invention may be natural proteins derived from any human or animal cells or tissues, chemically synthesized proteins, or proteins obtained by genetic recombination techniques. The protein may or may not be subjected to post-translational modifications such as sugar chain addition or phosphorylation.

The present invention also includes a polynucleotide encoding the above protein of the present invention. Examples of nucleotide sequences encoding a protein consisting of an amino acid sequence of any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180 include nucleotide sequences of any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179. The DNA includes cDNA, genomic DNA, and chemically synthesized DNA. In accordance with the degeneracy of the genetic code, at least one nucleotide in the nucleotide sequence encoding a protein consisting of an amino acid sequence of any one of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127,

That is, "stringent conditions" refer to overnight incubation at 37°C in a hybridization solution containing 30% formamide, 5 x SSC (0.75 M NaCl, 75mM trisodium citrate), 5 x Denhardt's solution, 0.5% SDS, 100 μ g/ml denatured, sheared salmon sperm DNA) followed by washing (three times) in 2 x SSC, 0.1% SDS for 10 minutes at room temperature, then followed by washing (two times) in 0.2 x SSC, 0.1% SDS for 10 minutes at 37°C (low stringency). Preferred stringent conditions are overnight incubation at 42 °C in a hybridization solution containing 40% formamide, followed by washing (three times) in 2 x SSC, 0.1% SDS for 10 minutes at room temperature, then followed by washing (two times) in 0.2 x SSC, 0.1% SDS for 10 minutes at 42°C (moderate stringency). More preferred stringent conditions are overnight incubation at 42°C in a hybridization solution containing 50% formamide, followed by washing (three times) in 2 x SSC, 0.1% SDS for 10 minutes at room temperature, followed by washing (two times) in 0.2 x SSC, 0.1% SDS for 10 minutes at 50°C (high stringency). The DNA sequence thus obtained must encode a protein capable of

activating NF- κ B.

The present invention also includes a polynucleotide comprising a nucleotide sequence which encodes a protein capable of activating NF- κ B and has a high sequence similarity to the nucleotide sequence of the polynucleotide according to above item (3) or (4). Typically these nucleotide sequence are 95% identical, preferably 97% identical, more preferably 98-99% identical, most preferably at least 99% identical to the nucleotide sequence of the polynucleotide according to above item (3) or (4) over the entire length thereof.

The above nucleotide sequence of the present invention can be used to produce the above protein using recombinant DNA techniques. In general, the DNA and peptide of the present invention can be obtained by:

- (A) cloning the DNA encoding the protein of the present invention;
- (B) inserting the DNA encoding the entire coding region of the protein or a part thereof into an expression vector to construct a recombinant vector;
- (C) transforming host cells with the recombinant vector thus constructed; and
- (D) culturing the obtained cells to express the protein or its analogue, and then purifying it by column chromatography.

General procedures necessary to handle DNA and recombinant host cells (e.g., *E. coli*) in the above steps are well known to those skilled in the art, and can be easily carried out in accordance with various laboratory manuals such as T. Maniatis et al., *supra*. All the enzymes, reagents, etc., used in these procedures are commercially available, and unless otherwise stated, such commercially available products can be used according to the use conditions specified by the manufacturer's instructions to attain completely its objects. The above steps (A) to (D) can be further illustrated in more details as follows.

Techniques for cloning the DNA encoding the protein of the present invention include, in addition to the methods described in the specification of the present application, PCR amplification using a synthetic DNA having a part of the nucleotide sequence of the present invention (e.g., any one of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120,

Thus, the present invention also includes a recombinant vector, which comprises the above DNA sequence. The expression vector for the protein of the present invention can be produced, for example, by excising the desired DNA fragment from the DNA encoding the protein of the present invention, and ligating the DNA fragment downstream of a promoter in a suitable expression vector.

Expression vectors for use in the present invention may be any vectors derived from prokaryotes (e.g., *E. coli*), yeast, fungi, insect viruses and vertebrate viruses so long as such vectors are replicable. However, the vectors should be selected to be compatible with

microorganisms or cells used as hosts. Suitable combinations of host cell – expression vector systems are selected depending on the desired expression product.

When microorganisms are used as hosts, plasmid vectors compatible with these microorganisms are generally used as replicable expression vectors for recombinant DNA molecules.

For example, the plasmids pBR322 and pBR327 can be used to transform *E. coli*. Plasmid vectors normally contain an origin of replication, a promoter, and a marker gene conferring upon a recombinant DNA a phenotype useful for selecting the cells transformed with the recombinant DNA. Example of such promoters include a β -lactamase promoter, lactose promoter and tryptophan promoter. Examples of such marker genes include an ampicillin resistance gene, and a tetracycline resistance gene. Examples of suitable expression vectors include the plasmids pUC18 and pUC19 in addition to pBR322, pBR327.

In order to express the DNA of the present invention in yeast, for example, YEp24 can be used as a replicable vector. The plasmid YEp24 contains the URA3 gene, which can be employed as a marker gene. Examples of promoters in expression vectors for yeast cells include promoters derived from genes for 3-phosphoglycerate kinase, glyceraldehyde-3-phosphate dehydrogenase and alcohol dehydrogenase.

Examples of promoters and terminators for use in expression vectors to express the DNA of the present invention in fungal cells include promoters and terminators derived from genes for phosphoglycerate kinase (PGK), glyceraldehyde-3-phosphate dehydrogenase (GAPD) and actin. Examples of suitable expression vectors include the plasmids pPGACY2 and pBSFAHY83.

Examples of promoters for use in expression vectors to express the DNA of the present invention in insect cells include a polyhedrin promoter and P10 promoter.

Recombinant vectors used to express the DNA of the present invention in animal cells normally contain functional sequences to regulate genes, such as an origin of replication, a promoter to be placed upstream of the DNA of the present invention, a ribosome-binding site, a polyadenylation site and a transcription termination sequence. Such functional sequences, which can be used to express the DNA of the present invention in eukaryotic cells, can be

Microorganisms or cells transformed with the replicable recombinant vector of the present invention can be selected from remaining untransformed parent cells based on at least one phenotype conferred by the recombinant vector. Phenotypes can be conferred by inserting at least one marker gene into the recombinant vector. Marker genes naturally contained in replicable vectors can be employed. Examples of marker genes include drug resistance genes such as neomycin resistance genes, and genes encoding dihydrofolate reductase.

25

Aspergillus nidulans, and *Acremonium chrysogenum* (ATCC 11550).

As insect cells, for example, *Spodoptera frugiperda* (Sf cells), High Five™ cells derived from eggs of *Trichoplusia*, etc., can be used when the virus is AcNPV. Examples of such animal cells include HEK 293 cells, COS-1 cells, COS-7 cells, Hela cells, and Chinese hamster ovary (CHO) cells. Among them, CHO cells and HEK 293 cells are preferred.

When cells are used as hosts, combinations of expression vectors and host cells to be used vary with experimental objects. According to such combinations, two types of expression (i.e. transient expression and constitutive expression) can be included.

"Transformation" of microorganisms and cells in the above step (C) refers to introducing DNA into microorganisms or cells by forcible methods or phagocytosis of cells and then transiently or constitutively expressing the trait of the DNA in a plasmid or an intra-chromosome integrated form. Those skilled in the art can carry out transformation by known methods [see e.g., "Idenshi Kougaku Handbook (Genetic Engineering Handbook)", an extra issue of "Jikken Igaku (Experimental Medicine)", YODOSHA CO., LTD.]. For example, in the case of animal cells, DNA can be introduced into cells by known methods such as DEAE-dextran method, calcium-phosphate-mediated transfection, electroporation, lipofection, etc. For stable expression of the protein of the present invention using animal cells, there is a method in which selection can be carried out by clonal selection of the animal cells containing the chromosomes into which the introduced expression vectors have been integrated. For example, transformants can be selected using the above selectable marker as an indication of successful transformation. In addition, the animal cells thus obtained using the selectable marker can be subjected to repeated clonal selection to obtain stable animal cell strains highly capable of expressing the protein of the present invention. When a dihydrofolate reductase (DHFR) gene is used as a selectable marker, one can culture animal cells while gradually increasing the concentration of methotrexate (MTX) and select the resistant strains, thereby amplifying the DNA encoding the protein of the present invention together with the DHFR gene to obtain animal cell strains having higher levels of expression.

The above transformed cells can be cultured under conditions which permit the expression of the DNA encoding the protein of the present invention to produce and

accumulate the protein of the present invention. In this manner, the protein of the present invention can be produced. Thus, the present invention also includes a process for producing a protein, which comprises culturing a transformed cell comprising the isolated polynucleotide according to above item (3) to (6) under conditions providing expression of the encoded protein and recovering the protein from the culture.

The above transformed cells can be cultured by methods known to those skilled in the art (see e.g., "Bio Manual Series 4", YODOSHA CO., LTD.). For example, animal cells can be cultured by various known animal cell culture methods including attachment culture such as Petri dish culture, multitray type culture and module culture, attachment culture in which cells are attached to cell culture carriers (microcarriers), suspension culture in which productive cells themselves are suspended. Examples of media for use in the culture include media commonly used for animal cell culture, such as D-MEM and RPMI 1640.

In order to separate and purify the protein of the present invention from the above culture, suitable combinations of per se known separation and purification methods can be used. Examples such methods include methods based on solubility, such as salting-out and solvent precipitation; methods based on the difference in charges, such as ion-exchange chromatography; methods mainly based on the difference in molecular weights, such as dialysis, ultrafiltration, gel filtration and SDS-polyacrylamide gel electrophoresis; methods based on specific affinity, such as affinity chromatography; methods based on the difference in hydrophobicity, such as reverse phase high performance liquid chromatography; and methods based on the difference in isoelectric points, such as isoelectric focusing. For example, a protein of the present invention can be recovered and purified from recombinant cell cultures by well-known methods including ammonium sulfate or ethanol precipitation, acid extraction, anion or cation exchange chromatography, phosphocellulose chromatography, hydrophobic interaction chromatography, affinity chromatography, hydroxyapatite chromatography and lectin chromatography. Most preferably, high performance liquid chromatography is employed for purification. Well known techniques for refolding proteins may be employed to regenerate active conformation when the polypeptide is denatured during intracellular synthesis, isolation or purification.

The protein of the present invention can also be produced as a fusion protein with another protein. These fusion proteins are also included within the present invention. For the expression of such fusion proteins, any vectors can be used so long as the DNA encoding the protein can be inserted into the vectors and the vectors can express the fusion protein. Examples of proteins to which a polypeptide of the present invention can be fused include glutathione S-transferase (GST) and a hexa-histidine sequence (6 x His). The fusion protein of the protein of the present invention with another protein can be advantageously purified by affinity chromatography using a substance with an affinity for the fusion partner protein. For example, fusion proteins with GST can be purified by affinity chromatography using glutathione as a ligand.

The present invention also includes an inhibitory protein, i.e., a protein capable of inhibiting the activity of the protein of above item (7). Examples of such inhibitory proteins include antibodies, or other proteins that bind to active sites of a protein of the above item (7), thereby inhibiting the expression of their activity.

The present invention also relates to an antibody that specifically binds the protein of the present invention or a fragment thereof, and to production of such an antibody. The antibody is not specifically limited so long as it can recognize the protein of the present invention. Examples of such antibodies include polyclonal antibodies, monoclonal antibodies and their fragments, single chain antibodies and humanized antibodies. Antibody fragments can be produced by known techniques. Examples of such antibody fragments include, but not limited to, F(ab')₂ fragments, Fab' fragments, Fab fragments and Fv fragments. The antibody that specifically binds the protein of the present invention can be produced using the protein of the present invention or a peptide thereof as an immunogen according to per se known process for producing antibodies or antisera. For example, a monoclonal or polyclonal antibody can be produced by administering the protein according to above item (1) or (2) as an antigen or epitope-bearing fragments to a non-human animal. Such methods are described, for example, in "Shin Idenshi Kougaku Handbook (New Genetic Engineering Handbook)", the third edition, an extra issue of "Jikken Igaku (Experimental Medicine)", YODOSHA CO., LTD.

When the resulting antibody is used to administer it to humans, it is preferably used as a humanized antibody or human antibody in order to reduce its immunogenicity. The humanized antibody can be produced using transgenic mice or other mammals. For a general review of these humanized antibodies and human antibodies, see, for example, Morrison, S.L. et al., Proc. Natl. Acad. Sci. USA, 81:6851-6855 (1984); Jones, P.T. et al., Nature 321:522-525 (1986); Hiroshi Noguchi, Igaku no Ayumi (J. Clin. Exp. Med.) 167:457-462 (1993); Takashi Matsumoto, Kagaku to Seibutsu (Chemistry and Biology) 36:448-456 (1998). Humanized chimeric antibodies can be produced by linking a V region of a mouse antibody to a C region of a human antibody. Humanized antibodies can be produced by substituting a sequence derived from a human antibody for a region other than a complementarity-determining region from a mouse monoclonal antibody. In addition, human antibodies can be directly produced in the same manner as the production of conventional monoclonal antibodies by immunizing the mice whose immune systems have been replaced with human immune systems. These antibodies can be used to isolate or to identify clones expressing the protein or to purify the

protein of the present invention from a cell extract or transformed cells producing the protein of the present invention. These proteins can also be used to construct ELISA, RIA (radioimmunoassay) and western blotting systems. These assay systems can be used for diagnostic purposes for detecting an amount of the protein of the present invention present in a body sample in a tissue or a fluid in the blood of an animal, preferably human. For example, they can be used for diagnosis of a disease characterized by undesirable activation of NF- κ B resulting from (expression) abnormality of the protein of the present invention, such as inflammation, autoimmune disease, infection (for example, HIV infection), bone disease, cancer and the like. In order to provide a basis for diagnosis of a disease, a standard value must be established. However, this is a well-known technique to those skilled in the art. For example, a method of calculating the standard value comprises binding a body fluid or a cell extract of normal individual of a human or an animal to an antibody against the protein of the present invention under a suitable condition for the complex formation, detecting the amount of the antibody-protein complex by chemical or physical means and then calculating the standard value for the normal sample using a standard curve prepared from a standard solution containing a known amount of an antigen (the protein of the present invention). The presence of a disease can be confirmed by deviation from the standard value obtained by comparison of the standard value with the value obtained from a sample of an individual latently suffering from a disease associated with the protein of the present invention. These antibodies can also be used as reagents for studying functions of the protein of the present invention.

The antibodies of the present invention can be purified and then administered to patients of a disease characterized by undesirable activation of NF- κ B resulting from (expression) abnormality of the protein of the present invention, such as inflammation, autoimmune disease, infection (such as HIV infection), bone disease, cancer and the like. Thus in another aspect, the present invention is a pharmaceutical composition which comprises the above antibody as an active ingredient, and therapy using the antibody of the present invention. In such pharmaceutical compositions, the active ingredient may be combined with other therapeutically active ingredients or inactive ingredients (e.g.,

- (a) determining the presence or absence of a mutation in the nucleotide sequence encoding the protein according to any one of claims 1, 2 or 7, in the genome of the subject, and/or
- (b) analyzing the amount of expression of said protein in a sample derived from said subject, wherein a diagnosis of disease is preferably made when the amount of the protein expressed is 2-fold or higher than normal, or half or lower than normal.

When the nucleotide sequence encoding the protein of above item (1), (2) or (7) which has a function of activating NF- κ B, contains a mutation according to the above step (a), the mutation may cause a disease associated with NF- κ B activation. When the amount of the expression of the protein of above item (1), (2) or (7) is different from the normal value according to the above step (b), the abnormal expression of the novel protein of the present invention which acts to activate NF- κ B may be responsible for diseases associated with NF- κ B activation. In the above step (a), determination of the presence or absence of a mutation in the nucleotide sequence of a the gene encoding the protein of above item (1), (2) or (7)

which has a function of activating NF- κ B, may involve RT-PCR using a part of the nucleotide sequences of genes encoding these proteins as a primer, followed by conventional DNA sequencing to detect the presence or absence of the mutation. PCR-SSCP [Genomics 5:874-879 (1989); "Shin Idenshi Kougaku Handbook (New Genetic Engineering Handbook)", an extra issue of "Jikken Igaku (Experimental Medicine)", YODOSHA CO., LTD.] can also be used to determine the presence or absence of the mutation.

Measurement of the amount of the expression of the protein in the above step (b) may involve, for example, using the antibody of above item (16).

The present invention also relates to a method for screening compounds which inhibit or promote NF- κ B activation using the proteins of the invention, which comprises the steps of:

- (a) providing a cell with a gene encoding a protein that activates NF- κ B, and a component that provides a detectable signal upon activation of NF- κ B;
- (b) culturing the transformed cell under conditions, which permit the expression of the gene in the transformed cell;
- (c) contacting the transformed cell with one or more compounds; and
- (d) measuring the detectable signal; and
- (e) isolating or identifying an activator compound and/or an inhibitor compound by measuring the detectable signal.

Further, it is preferable to isolate or identify as an activator compound, a compound that increases said detectable signal 2-fold or higher than normal, and to isolate or identify as an inhibitor compound, a compound that decreases said detectable signal half or less than normal.

Examples of components capable of providing a detectable signal include reporter genes. Reporter genes are used instead of directly detecting the activation of transcription factors of interest. The transcriptional activity of a promoter of a gene is analyzed by linking the promoter to a reporter gene and measuring the activity of the product of the reporter gene ("Bio Manual Series 4" (1994), YODOSHA CO., LTD.).

Any peptide or protein can be used so long as those skilled in the art can measure

the activity or amount of the expression product (including the amount of the produced mRNA) of the reporter genes. For example, enzymatic activity of chloramphenicol acetyltransferase, β -galactosidase, luciferase, etc., can be measured. Any reporter plasmids can be used to evaluate NF- κ B activation so long as the reporter plasmids have an NF- κ B recognition sequence inserted upstream of the reporter gene. For example, pNF- κ B-Luc (STRATAGEGE) can be used. Other examples include NF- κ B dependent reporter plasmids described in Tanaka S. et al., J. Vet. Med. Sci. Vol.59 (7); Rothe M. et al., Science Vol.269, p.1424-1427 (1995).

Any host cells may be used so long as NF- κ B activation can be detected in the host cells. Preferred host cells are mammalian cells such as 293-EBNA cells. Transformation and culture of the cells can be carried out as described above.

In a specific embodiment, the method for screening a compound which inhibits or promotes NF- κ B activation comprises culturing the transformed cell for a certain period of time, adding a certain amount of a test compound, measuring the reporter activity expressed by the cell after a certain period of time, and comparing the activity with that of a cell to which the test compound has not been added. The reporter activity can be measured by methods known in the art (see e.g., "Bio Manual Series 4" (1994), YODOSHA CO., LTD.). Examples of test compounds include, but not limited to, low molecular weight compounds and peptides. Test compounds may be artificially synthesized compounds or naturally occurring compounds. Test compounds may be a single compound or mixtures. Examples of such detectable signals which may be measured include the amount of mRNA or proteins for genes whose expression is known to be induced by NF- κ B activation (e.g., genes for IL-1 and TNF- α) in addition to the above reporter genes. The amount of mRNA can be measured, for example, by northern hybridization, RT-PCR, etc. The amount of proteins can be measured, for example, by using antibodies. The antibodies may be produced by known methods. Commercially available antibodies (from, e.g., Wako Pure Chemical Industries, Ltd.) can also be used.

It is also possible to produce a pharmaceutical composition according to the following steps (a) to (f):

(a) providing a cell with a gene encoding a protein that activates NF- κ B, and a component

that provides a detectable signal upon activation of NF- κ B;

- (b) culturing the transformed cell under conditions, which permit the expression of the gene in the transformed cell;
- (c) contacting the transformed cell with one or more candidate compounds;
- (d) measuring the detectable signal;
- (e) isolating or identifying an activator compound and/or an inhibitor compound by measuring the detectable signal; and
- (f) optimizing the isolated or identified compound as a pharmaceutical composition.

Further, it is preferable to isolate or identify as an activator compound, a compound that increases said detectable signal 2-fold or higher than normal, and to isolate or identify as an inhibitor compound, a compound that decreases said detectable signal half or less than normal.

The protein of the present invention may also be used in a method for the structure-based design of an agonist, antagonist or inhibitor of the protein, by:

- (a) determining in the first instance the three-dimensional structure of the protein;
- (b) deducing the three-dimensional structure for the likely reactive or binding site(s) of an agonist, antagonist or inhibitor;
- (c) synthesising candidate compounds that are predicted to bind to or react with the deduced binding or reactive site; and
- (d) testing whether the candidate compounds are indeed agonists, antagonists or inhibitor.

The present invention also includes a compound obtainable by the above screening method. However, the screening method of the present invention is not limited to the above method. The present invention also includes a process for producing the pharmaceutical composition by the method of above item (14).

There is no special limitation to the above candidate compounds. Such compounds include low molecular weight compounds and peptides. They may be artificially synthesised compounds and naturally occurring compounds. As the compounds obtained by the above screening methods have a function of inhibiting or promoting NF- κ B activation, they are useful as therapeutic or preventive pharmaceuticals for the treatment of diseases resulting from

A kit comprising at least one of (a) to (d) is useful for diagnosing a disease or susceptibility to a disease such as inflammation, autoimmune diseases, infectious diseases (e.g., HIV infection) and cancers.

Because NF- κ B is involved in a wide variety of pathological conditions such as inflammation, autoimmune diseases, cancers and viral infections, it is an attractive target for drug design and therapeutic intervention. Many experiments show that NF- κ B activity may have significant physiological effects [see e.g., *Ann. Rheum. Dis.* 57:738-741 (1998); *American Journal of Pathology* 152:793-803 (1998); *ARTHRITIS & RHEUMATISM* 40:226-236 (1997); *Am. J. Respir. Crit. Care Med.* 158:1585-1592 (1998); *J. Exp. Med.* 188:1739-1750 (1998); *Gut* 42:477-484 (1998); *The Journal of Immunology* 161:4572-4582 (1998); *Nature Medicine* 3:894-899 (1997)]. The finding of the new protein described herein capable of activating NF- κ B has provided a new method for controlling an abnormal NF- κ B function. Thus, the present invention also relates to use of a compound which inhibits the function of the protein capable of activating NF- κ B described above, for inhibiting NF- κ B activation. Further, the present invention relates to a method of using a compound which

activates the function of the protein capable of activating NF- κ B described above, for promoting NF- κ B activation. The compound obtained by the above screening method, which inhibits NF- κ B activation, is useful as a medicament to treat or prevent diseases characterized by undesirable activation of NF- κ B, such as inflammation, autoimmune diseases (such as rheumatoid arthritis, systemic lupus erythematosus, asthma, etc), infectious diseases, bone diseases, and graft rejection. Recently, it has also become apparent that NF- κ B activation controls apoptosis of cells. The compound obtained by the above screening method, which inhibits NF- κ B activation, may be capable of stimulating apoptosis. Diseases which may be treated by the induction of apoptosis include tumors.

Further, examples of diseases related to abnormality in NF- κ B activation include AIDS (acquired immunodeficiency syndrome), neurodegenerative diseases (Alzheimer's disease, Parkinson's disease, amyotrophic lateral sclerosis, etc.) ischemic disorders (i.e. those caused by cardiac infarction, reperfusion injury, etc), myelogenesis incompetency syndrome (aplastic anemia, etc), skin diseases (Toxic epidermal necrolysis, etc), proliferative nephritis (IgA nephritis, purpuric nephritis, lupus nephritis, etc) and fulminant hepatitis. Thus, a compound obtained by the above screening method, which inhibits or promotes NF- κ B activation, is useful as a medicament to treat or prevent these diseases.

In addition, the gene encoding the protein of the present invention is useful for gene therapy to treat various diseases such as cancers, autoimmune diseases, allergy diseases and inflammatory response. "Gene therapy" refers to administering into the human body a gene or a cell into which a gene has been introduced. The protein of the present invention and the DNA encoding the protein can also be used for diagnostic purposes.

The compound obtained by the screening method of the present invention or a salt thereof can be formulated into the above pharmaceutical compositions (e.g., tablets, capsules, elixirs, microcapsules, sterile solutions and suspensions) according to conventional procedures. The formulations thus obtained are safe and of low toxicity, and can be administered, for example, to humans and mammals (e.g., rats, rabbits, sheep, pigs, cattle, cats, dogs and monkeys). Administration to patients can be carried out by methods known in the art, such as intra-arterial injection, intravenous injection and subcutaneous injection. The dosage may

vary with the weight and age of the patient as well as a mode of administration, but those skilled in the art can appropriately select suitable dosages. When the compound can be encoded by DNA, the DNA can be inserted into a vector for gene therapy, and gene therapy can be carried out. The dosage and mode of administration may vary with the weight, age and symptoms of the patient, but those skilled in the art can appropriately select them. Thus, the present invention also relates to a pharmaceutical composition which comprises the above compound as an active ingredient.

In addition, the above compound is useful as a medicament to treat or prevent diseases characterized by abnormal NF- κ B activity, such as inflammation, autoimmune diseases, viral diseases, infectious diseases, cancers and bone diseases. Thus, the present invention also relates to a pharmaceutical composition for inflammation, autoimmune diseases, viral diseases, infectious diseases, cancers, bone diseases, etc., which comprises the above compound. Specifically, the compound is useful as a therapeutic and/or prophylactic drug against, for example, rheumatoid arthritis, osteoarthritis, systemic lupus erythematosus, diabetes, sepsis, asthma, allergic rhinitis, ischemic heart diseases, inflammatory intestinal diseases, subarachnoid hemorrhage, viral hepatitis, AIDS, atherosclerosis, atopic dermatosis, viral infections, Crohn's disease, diabetes, gout, hepatitis, multiple sclerosis, cardiac infarction, nephritis, osteoporosis, Alzheimer's, Parkinson's disease, Huntington's chorea, psoriasis, amyotrophic lateral sclerosis, or aplastic anemia.

The present invention also relates to the use of a pharmaceutical composition produced according to above item (14) for manufacturing a medicament against inflammation, autoimmune diseases, viral diseases, cancers, infectious diseases, bone diseases, etc. The present invention also includes an antisense oligonucleotide against a gene of any one of above items (3) to (6). An antisense oligonucleotide refers to an oligonucleotide complementary to the target gene sequence. The antisense oligonucleotide can inhibit the expression of the target gene by inhibiting RNA functions such as translation to proteins, transport to the cytoplasm and other activity necessary for overall biological functions. In this case, the antisense oligonucleotide may be RNA or DNA. The DNA sequence of the present invention can be used to produce an antisense oligonucleotide capable of hybridizing with the mRNA

transcribed from the gene encoding the protein of the present invention. It is known that an antisense oligonucleotide generally has an inhibitory effect on the expression of the corresponding gene (see e.g., Saibou Kougaku Vol.13, No.4 (1994)). The oligonucleotide containing an antisense coding sequence against a gene encoding the protein of the present invention can be introduced into a cell by standard methods. The oligonucleotide effectively blocks the translation of mRNA of the gene encoding the protein of the present invention, thereby blocking its expression and inhibiting undesirable activity.

The oligonucleotide of the present invention may be a naturally occurring oligonucleotide or its modified form [see e.g., Murakami & Makino, Saibou Kougaku Vol.13, No.4, p.259-266 (1994); Akira Murakami, Tanpakushitsu Kakusan Kouso (PROTEIN, NUCLEIC ACID AND ENZYME) Vol.40, No.10, p.1364-1370 (1995), Tunenari Takeuchi et al., Jikken Igaku (Experimental Medicine) Vol. 14, No. 4 p85-95(1996)]. Thus, the oligonucleotide may have modified sugar moieties or inter-sugar moieties. Examples of such modified forms include phosphothioates and other sulfur-containing species used in the art. According to several preferred embodiments of the present invention, at least one phosphodiester bond in the oligonucleotide is substituted with the structure which can enhance the ability of the composition to permeate cellular regions where RNA with the activity to be regulated is located.

Such substitution preferably involves a phosphorothioate bond, a phosphoramidate bond, methylphosphonate bond, or a short-chain alkyl or cycloalkyl structure. The oligonucleotide may also contain at least some modified base forms. Thus, it may contain purine and pyrimidine derivatives other than naturally occurring purine and pyrimidine. Similarly, the furanosyl moieties of the nucleotide subunits can be modified so long as the essential purpose of the present invention is attained. Examples of such modifications include 2'-O-alkyl and 2'-halogen substituted nucleotides. Examples of modifications in sugar moieties at their 2-position include OH, SH, SCH₃, OCH₃, OCN or O(CH₂)_nCH₃, wherein n is 1 to about 10, and other substituents having similar properties. All the analogues are included in the scope of the present invention so long as they can hybridize with the mRNA of the gene of the present invention to inhibit functions of the mRNA.

The oligonucleotide of the present invention contains about 3 to about 50 nucleotides, preferably about 8 to about 30 nucleotides, more preferably about 12 to about 25 nucleotides. The oligonucleotide of the present invention can be produced by the well-known solid phase synthesis technique. Devices for such synthesis are commercially available from some manufactures including Applied Biosystems. Other oligonucleotides such as phosphothioates can also be produced by methods known in the art.

The oligonucleotide of the present invention is designed to hybridize with the mRNA transcribed from the gene of the present invention. Those skilled in the art can easily design an antisense oligonucleotides based on a given gene sequence (For example, Murakami and Makino: Saibou Kougaku Vol. 13 No.4 p259-266 (1994), Akira Murakami: Tanpakushitsu Kakusan Kouso (PROTEIN, NUCLEIC ACID AND ENZYME) Vol. 40 No.10 p1364-1370 (1995), Tunenari Takeuchi et al., Jikken Igaku (Experimental Medicine) Vol. 14 No. 4 p85-95 (1996)). Recent sutudy suggests that antisense oligonucleotides which are designed in a region containing 5' region of mRNA, preferably,the translation initaiation site,are most effective for the inhibition of the expression of a gene.The length of the antisense oligonucleotides is preferably 15 to 30 nucleotides and more preferably 20 to 25 nucleotides. It is important to confirm no interaction with other mRNA and no formation of secondary structure in the oligonucleotide sequence by homology search. The evaluation of whether the designed antisense oligonucleotide is functional or not can be determined by introducing the antisense oligonucleotide into a suitable cell and measuring the amount of the target mRNA, for example by northern blotting or RT-PCR, or the amount of the target protein, for example by western blotting or fluorescent antibody technique, to confirm the effect of expression inhibition

Another method includes the triple helix technique. This technique involves forming a triple helix on the targeted intra-nuclear DNA sequence, thereby regulating its gene expression, mainly at the transcription stage. The oligonucleotide is designed mainly in the gene region involved in the transcription and inhibits the transcription and the production of the protein of the present invention. Such RNA, DNA and oligonucleotide can be produced using known synthesizers.

The oligonucleotide may be introduced into the cells containing the target nucleic acid sequence by any of DNA transfection methods such as calcium phosphate method, electroporation, lipofection, microinjection, or gene transfer methods including the use of gene transfer vectors such as viruses. An antisense oligonucleotide expression vector can be prepared using a suitable retrovirus vector, then the expression vector can be introduced into the cells containing the target nucleic acid sequence by contacting the vector with the cells in vivo or ex vivo.

The DNA of the present invention can be used in the antisense RNA/DNA technique or the triple helix technique to inhibit NF- κ B activation mediated by the protein of the present invention.

The antisense oligonucleotide against the gene encoding the protein of the present invention is useful as a medicament to treat or prevent diseases characterized by undesirable activation of NF- κ B, such as inflammation, autoimmune diseases, infectious diseases (e.g., HIV infection) and cancers. Thus, the present invention also includes a pharmaceutical composition which comprises the above antisense oligonucleotide as an active ingredient. The antisense oligonucleotide can also be used to detect such diseases using northern hybridization or PCR.

The present invention also includes a ribozyme which inhibits NF- κ B activation. A ribozyme is an RNA capable of recognizing a nucleotide sequence of a nucleic acid and cleaving the nucleic acid (see e.g., Hiroshi Yanagawa, "Jikken Igaku (Experimental Medicine) Bioscience 12: New Age of RNA). The ribozyme can be produced so that it cleaves the selected target RNA (e.g., mRNA encoding the protein of the present invention). Based on the nucleotide sequence of the DNA encoding the protein of the present invention, the ribozyme specifically cleaving the mRNA of the protein of the present invention can be designed. Such ribozyme has a complementary sequence to the mRNA for the protein of the present invention, complementarily associates with the mRNA and then cleaves the mRNA, which results in reduction or entire loss of the expression of the protein of the present invention. The level of the reduction of the expression is dependent on the level of the ribozyme expression in the target cells.

There are two types of ribozyme commonly used: a hammerhead ribozyme and a hairpin ribozyme. In particular, hammerhead ribozymes have been well studied regarding their primary and secondary structure necessary for their cleavage activity, and those skilled in the art can easily design the ribozymes nucleotided solely on the nucleotide sequence information for the DNA encoding the protein of the present invention [see e.g., Iida et al., Saibou Kougaku Vol.16, No.3, p.438-445 (1997); Ohkawa & Taira, Jikken Igaku (Experimental Medicine) Vol.12, No.12, p.83-88 (1994)]. It is known that the hammerhead ribozymes have a structure consisting of two recognition sites (recognition site I and recognition site II forming a chain complementary to target RNA) and an active site, and cleave the target RNA at the 3'end of its sequence NUX (wherein N is A or G or C or U, and X is A or C or U)after the formation of a complementary pair with the target RNA in the recognition sites. In particular, the sequence GUC (or GUA) has been found to have the highest activity [see e.g., Koizumi, M. et al., Nucl. Acids Res. 17:7059-7071 (1989); Iida et al., Saibou Kougaku Vol.16, No.3, p.438-445 (1997); Ohkawa & Taira, Jikken Igaku (Experimental Medicine) Vol.12, No.12, p.83-88 (1994); Kawasaki & Taira, Jikken Igaku (Experimental Medicine) Vol.18, No.3, p.381-386 (2000)].

Therefore the sequence GTC (or GTA) is searched out, and a ribozyme is designed to form several, up to 10 to 20 complementary base pairs around that sequence. The suitability of the designed ribozyme can be evaluated by checking whether the prepared ribozyme can cleave the target mRNA in vitro according to the method described for example in Ohkawa & Taira, Jikken Igaku (Experimental Medicine) Vol.12, No.12, p.83-88 (1994). The ribozyme can be prepared by methods known in the art to synthesize RNA molecules.

Alternatively, the sequence of the ribozyme can be synthesized on a DNA synthesizer and inserted into various vectors containing a suitable RNA polymerase promoter (e.g., T7 or SP6) to enzymatically synthesize an RNA molecule in vitro. Such ribozymes can be introduced into cells by gene transfer methods such as microinjection. Another method involves inserting a ribozyme DNA into a suitable expression vector and introducing the vector into cell strains, cells or tissues. Suitable vectors can be used to introduce the ribozyme into a selected cell. Examples of vectors commonly used for such purpose include

Methods using a cDNA library containing a lot of non-full-length cDNAs are inefficient in obtaining many genes (cDNAs) having functions. Therefore libraries with a high ratio of the number of the full-length cDNA clones to the total number of the clones are necessary. "Full-length cDNA" refers to a complete DNA copy of mRNA from a gene. The cDNA libraries produced using the oligo-capping method contain full-length cDNA clones in a ratio of 50 to 80%, namely, a 5 to 10-fold increase in full-length cDNA clones compared to the cDNA libraries produced by prior art methods (Sumio Sugano, the monthly magazine BIO INDUSTRY Vol.16, No.11, p.19-26). Full-length cDNA clones are essential for protein expression in functional analyses of genes, and full-length cDNA clones themselves are very important materials for activity measurement. Thus, cloning of full-length cDNA is necessary for functional analyses of genes. Sequencing of the cDNA not only provides important information for establishing the primary sequence of the protein encoded by the cDNA, but also reveals the entire exon sequence. Thus, the full-length cDNA provides valuable information for identifying a gene, such as information for determining the primary sequence of a protein, exon-intron structure, the transcription initiation site of mRNA, the location of a promoter, etc.

44

The process of the present invention uses an in vitro system or a cell-based system, preferably a cell-based system. Examples of such cells include cells of prokaryotes such as *E. coli*, microorganisms such as yeast and fungi, as well as insects and animals. Preferred examples include animal cells, in particular, 293-EBNA cells and NIH3T3 cells.

45

simultaneously. One embodiment of the process of the present invention is detailed in Examples herein. Alternatively, a screening system for obtaining a gene capable of inhibiting NF- κ B activation can also be constructed by cotransfecting an expression vector comprising full-length cDNA and a reporter gene into cells, stimulating the cells with IL-1 or TNF- α and the like, and selecting a clone having subnormally increased reporter activity.

However, the process of the present invention is not limited to these embodiments.

Because the cDNA of the present invention is full-length, its 5' end sequence is the transcription initiation site of the corresponding mRNA. Therefore the cDNA sequence can be used to identify the promoter region of the gene by comparing the cDNA with the genomic nucleotide sequence. Genomic nucleotide sequences are available from various databases when the sequences have been deposited in the databases. Alternatively, the cDNA can also be used to clone the desired sequence from a genomic library, for example, by hybridization, and determine its nucleotide sequence. Thus, by comparing the nucleotide sequence of the cDNA of the present invention with a genomic sequence, the promoter region of the gene located upstream the cDNA can be identified. In addition, the promoter fragment thus identified can be used to construct a reporter plasmid for evaluating the expression of the gene. In general, the DNA fragment spanning 2kb (preferably 1kb) upstream from the transcription initiation site can be inserted upstream of the reporter gene to produce the reporter plasmid. The reporter plasmid can be used to screen for a compound which enhances or reduces the expression of the gene. For example, such screening can be carried out by transforming a suitable cell with the reporter plasmid, culturing the transformed cell for a certain period of time, adding a certain amount of a test compound, measuring the reporter activity expressed by the cell after a certain period of time, and comparing the activity with that of a cell to which the test compound has not been added. These methods are also included in the scope of the present invention.

The present invention also relates to a computer-readable medium on which a sequence data set has been stored, said sequence data set comprising at least one nucleotide sequence selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70,

72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179, and/or at least one amino acid sequence selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180.

In another aspect, the present invention relates to a method for calculating a homology, which comprises comparing data on the above medium with data of other nucleotide sequences. Thus, the gene and amino acid sequence of the present invention provide valuable information for determining their secondary and tertiary structure, e.g., information for identifying other sequence having a similar function and high homology. These sequences are stored on the computer-readable medium, then a database is searched using data stored in a known macromolecule structure program and a known search tool such as GCG. In this manner, a sequence in a database having a certain homology can be easily found.

The computer-readable medium may be any composition of materials used to store information or data. Examples of such media include commercially available floppy disks, tapes, chips, hard drives, compact disks and video disks. The data on the medium allows a method for calculating a homology by comparing the data with other nucleotide sequence data. This method comprises the steps of providing a first polynucleotide sequence containing the polynucleotide sequence of the present invention for the computer-readable medium, and then comparing the first polynucleotide sequence with at least one-second polynucleotide or polypeptide sequence to identify the homology.

The present invention also relates to an insoluble substrate to which polynucleotide comprising all or part of the nucleotide sequences selected from the group consisting of SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98,

100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179, are fixed. A plurality of the various polynucleotides which are DNA probes are fixed on a specifically processed solid substrate such as slide glass to form a DNA microarray and then a labeled target polynucleotide is hybridized with the fixed polynucleotides to detect a signal from each of the probes. The data obtained is analyzed and the gene expression is determined.

The present invention further relates to an insoluble substrate to which polypeptides comprising all or part of the amino acid sequences selected from the group consisting of SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87, 89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180, are fixed. By mixing organism-derived cell extract with the insoluble substrate on which these proteins are fixed, it is possible to isolate or identify cell-derived components such as proteins captured on the insoluble substrate that can be expected to be useful in diagnosis or drug development.

EXAMPLES

The following examples further illustrate, but do not limit the present invention.

Example 1: Construction of a full-length cDNA library using the oligo-capping method

(1) Preparation of RNA from human lung fibroblasts (Cryo NHLF)

Human lung fibroblasts (Cryo NHLF: purchased from Sanko Junyaku Co., Ltd.) were cultured according to the attached protocol. After repeating subculturing the cells to obtain fifty 10cm dishes containing the resulting culture, the cells were recovered with a cell scraper. Then, total RNA was obtained from the recovered cells by using the RNA extraction reagent ISOGEN (purchased from NIPPON GENE) according to the manufacturer's protocol. Then, poly A⁺ RNA was obtained from the total RNA by using an oligo-dT cellulose column according to Maniatis et al., supra.

(2) Preparation of RNA from mouse ATDC5 cells

ATDC5, a cell strain cloned from mouse EC (embryonal carcinoma) (Atsumi, T. et al.: Cell Diff. Dev., 30: p109-116)(1990) was repeatedly subcultured to obtain fifty 10cm dishes containing the resultant culture. Thereafter, poly A⁺ RNA was obtained by a method similar to that of (1) above. Culture of ATDC5 cells was performed according to the method described in Atsumi, T. et al.: Cell Diff. Dev., 30: p109-116 (1990).

(3) Construction of a full-length cDNA library by the oligo-capping method

A full-length cDNA library was constructed from poly A⁺ RNA of the above human lung fibroblasts and ATDC5 cells by the oligo-capping method according to the method of Sugano S. et al. [e.g., Maruyama, K. & Sugano, S., Gene, 138:171-174 (1994); Suzuki, Y. et al., Gene, 200:149-156 (1997); Suzuki, Y. & Sugano, S. "Shin Idenshi Kougaku Handbook (New Genetic Engineering Handbook)", the third edition (1999), an extra issue of "Jikken Igaku (Experimental Medicine)", YODOSHA CO., LTD.].

(4) Preparation of plasmid DNA

The full-length cDNA library constructed as above was transfected into E. coli strain TOP 10 by electroporation, then spread on LB agar medium containing 100 μ g/ml of ampicillin, and incubated overnight at 37°C. Then, using QIAwell 96 Ultra Plasmid Kit (QIAGEN) according to the manufacturer's protocol, the plasmids were recovered from the colonies grown on ampicillin-containing LB agar medium.

Example 2: Cloning of DNA capable of activating NF- κ B

(1) Screening of the cDNA encoding the protein capable of activating NF- κ B

293-EBNA cells (purchased from Invitrogen) were grown to 1×10^4 cells/100 μ l/well in a 96 well plate for cell culture for 24 hours at 37°C (in the presence of 5% CO₂) using 5% FBS containing DMEM medium. Then, 50ng of pNF κ B-Luc (purchased from STRATAGENE) and 2 μ l of the full-length cDNA expression vector prepared in above Example 1.(4) were cotransfected into the cells in a well using FuGENE 6 (purchased from Roche) according to the manufacturer's protocol. After 24 hours of culture at 37°C, the reporter activity of NF- κ B (luciferase activity) was measured using long-term luciferase assay system, PIKKA GENE LT2.0 (TOYO INK) according to the attached manufacturer's

activity was measured using Wallac ARVOTMST 1420 MULTILABEL COUNTER (Perkin Elmer).

(2) DNA sequencing

The above screening was carried out for 155,000 clones, and plasmids showing a 5-fold or more increase in luciferase activity compared to that of the control experiment (luciferase activity of the cell into which vacant vector pME18S-FL3 is introduced instead of full-length cDNA expression vector) were selected. One pass sequencing was carried out from the 5' end of the cloned cDNA (sequencing primer: 5'-CTTCTGCTCTAAAAGCTGCG-3' (SEQ ID NO: 181)) and from the 3' end (sequencing primer: 5'-CGACCTGCAGCTCGAGCACA-3' (SEQ ID NO: 182)) so that as long sequence as possible is determined. The sequencing was carried out using the reagent Thermo Sequenase II Dye Terminator Cycle Sequencing Kit (Amersham Pharmacia Biotech) or BigDye Terminator Cycle Sequencing FS Ready Reaction Kit (Applied Biosystems) and the device ABI PRISM 377 sequencer or ABI PRISM 3100 sequencer according to the manufacturer's instructions.

(3) Database analysis of the obtained clones

BLAST (Basic local alignment search tool) searching [S. F. Altschul et al., J. Mol. Biol., 215:403-410 (1990)] was carried out in GenBank for the obtained nucleotide sequence. The results showed that 148 clones represented 90 genes encoding new proteins capable of activating NF- κ B.

(4) Full-length sequencing

The full-length DNA sequences for the 90 new clones were determined (SEQ ID NOS: 2, 4, 6, 8, 10, 12, 14, 16, 18, 20, 22, 24, 26, 28, 30, 32, 34, 36, 38, 40, 42, 44, 46, 48, 50, 52, 54, 56, 58, 60, 62, 64, 66, 68, 70, 72, 74, 76, 78, 80, 82, 84, 86, 88, 90, 92, 94, 96, 98, 100, 102, 104, 106, 108, 110, 112, 114, 116, 118, 120, 122, 124, 126, 128, 130, 132, 134, 136, 138, 140, 142, 144, 146, 148, 150, 151, 153, 155, 157, 159, 161, 163, 165, 167, 169, 171, 173, 175, 177 and 179). The amino acid sequences of the protein coding regions (open reading frames) were deduced (SEQ ID NOS: 1, 3, 5, 7, 9, 11, 13, 15, 17, 19, 21, 23, 25, 27, 29, 31, 33, 35, 37, 39, 41, 43, 45, 47, 49, 51, 53, 55, 57, 59, 61, 63, 65, 67, 69, 71, 73, 75, 77, 79, 81, 83, 85, 87,

89, 91, 93, 95, 97, 99, 101, 103, 105, 107, 109, 111, 113, 115, 117, 119, 121, 123, 125, 127, 129, 131, 133, 135, 137, 139, 141, 143, 145, 147, 149, 152, 154, 156, 158, 160, 162, 164, 166, 168, 170, 172, 174, 176, 178 and 180).

Example 3: Screening compounds inhibiting NF- κ B activation

293-EBNA cells were seeded on 5% FBS containing DMEM medium in a 96-well cell culture plate to a final cell density of 1×10^4 cells/100 μ l/well, and cultured for 24 hours at 37°C in the presence of 5% CO₂. Then, 50ng of the expression vector comprising the gene encoding NF- κ B activating protein of SEQ ID NO: 5, 9, 17, 21, 35, 37, 41, 53, 57, 63, 67, 71, 75, 81, 87, 91, 93, 97, 121, 123, 129, 154, 158, 162, 168, 170, 172, 176 or 178, and 50ng of the reporter plasmid pNF κ B-Luc were cotransfected into the cells in a well using FuGENE 6. After 1 hour, the proteasome inhibitor MG-132 (purchased from CALBIOCHEM) (Uehara T. et al., J. Biol. Chem. 274, p.15875-15882 (1999); Wang X. C. et al., Invest. Ophthalmol. Vis. Sci. 40, p.477-486) was added to the culture to final concentrations of 0.1 μ M, 0.5 μ M, 1.0 μ M and 10 μ M, respectively. After 24 hours of culture at 37°C, the reporter activity was measured using PIKKA GENE LT2.0. The results showed that MG132 inhibited the expression of the reporter gene (Fig. 1 to Fig. 29).

INDUSTRIAL APPLICABILITY

As described above, the present invention provides industrially highly useful proteins capable of activating NF- κ B and genes encoding the proteins. The proteins of the present invention and the genes encoding the proteins allow not only screening for compounds useful for treating and preventing diseases associated with the excessive activation or inhibition of NF- κ B, but also production of diagnostics for such diseases. The genes of the present invention are also useful as a gene source used for gene therapy.

All publications, patents and patent applications cited herein are incorporated herein in their entirety.